Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 60–64, January 1996
Biophysics

A mean field model of ligand–protein interactions: Implications for the structural assessment of human immunodeficiency virus type 1 protease complexes and receptor-specific binding

GENNADY M. VERKHIVKER AND PAUL A. REJTO

Agouron Pharmaceuticals Inc., 3565 General Atomics Court, San Diego, CA 92121

Communicated by Hans C. Andersen, Stanford University, Stanford, CA, September 15, 1995 (received for review March 29, 1995)

ABSTRACT **We propose a general mean field model of ligand–protein interactions to determine the thermodynamic equilibrium of a system at finite temperature. The method is employed in structural assessments of two human immunodeficiency virus type 1 protease complexes where the gross effects of protein flexibility are incorporated by utilizing a data base of crystal structures. Analysis of the energy spectra for these complexes has revealed that structural and thermodynamic aspects of molecular recognition can be rationalized on the basis of the extent of frustration in the binding energy landscape. In particular, the relationship between receptor-specific binding of these ligands to human immunodeficiency virus type 1 protease and a minimal frustration principle is analyzed.**

Understanding the principles of molecular recognition is a long-standing problem of formidable complexity in molecular biology. Theoretical approaches to assess binding affinity and receptor specificity of a designed molecule prior to its synthesis are in high demand by the newly emerging interdisciplinary field of structure-based drug design (1–7). A necessary precursor for these studies is prediction of the structure of ligand–protein complexes. This problem has been approached by a number of stochastic optimization methods (8–19), and docking of a flexible ligand to a protein with a rigid backbone and flexible side chains, starting from their unbound forms, is now feasible (16, 18). However, current structural assessments typically are limited to docking either rigid or flexible ligands to a rigid protein in its bound conformation. This simplified scenario of molecular recognition still entails an enormous number of configurations that must be searched rapidly and consistently to determine the global free-energy minimum, and represents another manifestation of the Levinthal paradox (20).

Ambiguous structural predictions from docking simulations, when "false" solutions are ranked with energies similar to those of native structures, are believed to be caused by an incomplete energy function and a rough energy landscape (12, 14, 18). The free-energy minimum can be determined exactly on a lattice for short polypeptides that have no long-range internal interactions and are bound to an undeformed receptor surface that produces a fixed molecular field (21, 22). Critically, this model simplifies the complex energy landscape by removing the competition between short and long-range interactions. The frustration of the system and the roughness of the landscape (23, 24) are thereby alleviated, and the global solution can be determined unambiguously. Lattice models of protein folding (23–34) have revealed that minimal frustration is a critical prerequisite for fast kinetics and thermodynamic stability of proteins, and that stable proteins satisfy the principle of minimal frustration (23, 24, 29, 32–37). Ligand–protein complexes are also frustrated systems that typically have a multitude of low-energy metastable states and a rough energy landscape arising from competition between the internal constraints and the intermolecular ligand–protein interactions. Therefore, predicting the structure of a ligand–protein complex may be difficult not only because of the enormous number of accessible conformations but also due to the complex and frustrated character of the ligand–protein interactions on the underlying potential energy surface.

We formulate a mean field model of ligand–protein interactions to assess the structure of ligand–protein complexes at finite temperature. Mean field approaches are useful for examining complex interactions in multiparticle systems, such as structure prediction of small proteins within a given topological pattern (38), intramolecular conformational optimization (39), side-chain placement in proteins (40, 41), and prediction of mutant energetics (42). In general, self-consistent free-energy minimization based on a mean field approximation may lead to metastable solutions that depend on the initial conditions. When the energy landscape is frustrated by structurally dissimilar metastable states, mean field theory may fail to predict consistently the global free-energy minimum. Consequently, sensitivity of the solution to initial conditions is a signature of a frustrated energy landscape, and the self-consistent iteration procedure converges to a unique minimum only when the system is unfrustrated.

Mean field theories in protein structure prediction typically are valid only within a given topological pattern of the protein backbone (38, 43). In this application, we use a data base of 54 crystallographically determined human immunodeficiency virus type 1 (HIV-1) complexes (1–3) as a finite basis set of different protein conformations and assume that this data base spans a statistically representative set of structures that describe possible protein responses. While these complexes manifest the same general topological pattern of the protein structure in its bound conformation, substantial protein backbone and side-chain flexibility is retained (1). The motivation for using crystal structures is to effectively characterize the different protein conformational substates, which are inherently difficult to determine by molecular dynamics simulation because they are separated by large energy barriers. In this mean field model, the superimposed protein conformations describe dynamic fluctuations of a single protein at equilibrium, and, as in linear-response theory (44), these equilibrium fluctuations determine the nonequilibrium response of the protein to ligand binding.

To obtain the energy spectrum, the protein conformers in the database were first superimposed by least square superposition of the C^{α} atoms. Then, a given ligand was docked to each rigid protein conformer (45). During this procedure, which minimizes the energy of the ligand–protein complex by a simulated evolution approach (45, 46), the internal degrees of freedom of the ligand are held fixed. The spectrum includes both the energies of ligand–protein complexes that were optimized during the docking procedure and the interaction energies between "random" ligand–protein arrangements.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

These random complexes were obtained by retaining the ligand in its docked position while replacing the protein conformer to which it was docked with all other protein conformers in the crystal-structure data base. The model of ligand–protein interactions and the subsequent generation of the energy spectra are based on the mean field approximation that the protein responses do not depend on the ligand conformations. As stated previously, the conformations of the protein are obtained from a data base of possible protein states, independent of the conformation of the ligand, and random conformations of the complexes are generated from unrelated protein and ligand states.

Any complexity or frustration of the energy spectrum, which arises from competition between the native complex and random ligand–protein arrangements, is retained in this model. The energies of nonoptimized or random ligand–protein structures, which have little structural similarity to the native complex, form a continuous part of the energy spectrum. Consistent with both the energy gap (28) and the more general stability gap concepts (35–37), we measure the extent of frustration in the system by monitoring the energy difference between the onset of this continuum and the native complex rather than the difference between the energy of the native state and the next highest energy level (33).

The proposed model of ligand–protein interactions incorporates the gross effects of protein flexibility by using mean field theory with a finite number of protein conformations. To reproducibly determine the potential energy minimum of the ligand docked to each protein conformer, we neglect the effects of ligand flexibility in this study. Nevertheless, this model is general and can be used for the structural prediction of ligand–protein complexes given a data base of protein backbone conformations, a library of rotamer side-chain conformations, and flexible ligands, where the combinatorial explosion of possible states would demand a mean field approximation. We examine the energy spectra of HIV-1 complexes with two inhibitors that reveal salient aspects of the ligand–protein binding process. In particular, the relationship between the receptor-specific binding of ligands to HIV-1 protease relative to HIV-2 protease and the frustration of their energy landscapes is highlighted.

Mean Field Model

The Hamiltonian of a ligand–protein system may be expressed as the sum of the ligand H_{ligand} and protein H_{protein} intramolecular energies and the intermolecular ligand–protein interaction energy $H_{\text{ligand–protein interaction}}$ (hereafter designated $H_{\text{lig.-prot.}}$), namely

$$H(\mathbf{r}) = H_{\text{ligand}}(\mathbf{r}_{\text{ligand}}) + H_{\text{protein}}(\mathbf{r}_{\text{protein}}) + H_{\text{lig.-prot.}}(\mathbf{r}_{\text{ligand}}, \mathbf{r}_{\text{protein}}), \quad [1]$$

where $\mathbf{r} = (\mathbf{r}_{\text{ligand}}, \mathbf{r}_{\text{protein}})$.

The main idea of a general mean field model for determining the thermodynamically stable structure of a ligand–protein complex is to replace the exact forces acting on protein atoms from a ligand, and on ligand atoms from a protein, by the effective mean field. The crucial assumption of mean field theory is the factorization of the equilibrium probability distribution

$$p(\mathbf{r}_{\text{ligand}}, \mathbf{r}_{\text{protein}}) = p(\mathbf{r}_{\text{ligand}})p(\mathbf{r}_{\text{protein}}).$$

The first variation of the free energy δF when $p(\mathbf{r})$ is varied by the amount $\delta p(\mathbf{r})$ is given by

$$\delta F = \int \Phi(\mathbf{r}_{\text{ligand}})\delta p(\mathbf{r}_{\text{ligand}})d\mathbf{r}_{\text{ligand}} + \int \Psi(\mathbf{r}_{\text{protein}})\delta p(\mathbf{r}_{\text{protein}})d\mathbf{r}_{\text{protein}} + \int [H_{\text{ligand}}(\mathbf{r}_{\text{ligand}}) + H_{\text{protein}}(\mathbf{r}_{\text{protein}}) + kT \ln p(\mathbf{r})]\delta p(\mathbf{r})d\mathbf{r},$$

where we have defined the mean field potentials acting on the ligand coordinate $\mathbf{r}_{\text{ligand}}$ and the protein coordinate $\mathbf{r}_{\text{protein}}$ to be

$$\Phi(\mathbf{r}_{\text{ligand}}) = \int H_{\text{lig.-prot.}}(\mathbf{r}_{\text{ligand}}, \mathbf{r}_{\text{protein}})p(\mathbf{r}_{\text{protein}})d\mathbf{r}_{\text{protein}}$$

and

$$\Psi(\mathbf{r}_{\text{protein}}) = \int H_{\text{lig.-prot.}}(\mathbf{r}_{\text{ligand}}, \mathbf{r}_{\text{protein}})p(\mathbf{r}_{\text{ligand}})d\mathbf{r}_{\text{ligand}}.$$

Given the fact that $p(\mathbf{r}_{\text{ligand}})$ and $p(\mathbf{r}_{\text{protein}})$ are both normalized probability distributions, it follows that $\delta p(\mathbf{r}_{\text{ligand}}) = \int \delta p(\mathbf{r})d\mathbf{r}_{\text{protein}}$, and $\delta p(\mathbf{r}_{\text{protein}}) = \int \delta p(\mathbf{r})d\mathbf{r}_{\text{ligand}}$. When using this result, the condition that $\delta F = 0$ is valid for arbitrary $\delta p(\mathbf{r})$ leads to the following:

$$H_{\text{ligand}}(\mathbf{r}_{\text{ligand}}) + H_{\text{protein}}(\mathbf{r}_{\text{protein}}) + \Phi(\mathbf{r}_{\text{ligand}}) + \Psi(\mathbf{r}_{\text{protein}}) + kT \ln p(\mathbf{r}_{\text{ligand}}, \mathbf{r}_{\text{protein}}) = 0.$$

Consequently, the Hamiltonian $\hat{H}$ that approximates ligand–protein interactions of the original system by a mean field is given by

$$\hat{H}(\mathbf{r}) = H_{\text{ligand}}(\mathbf{r}_{\text{ligand}}) + H_{\text{protein}}(\mathbf{r}_{\text{protein}}) + \Phi(\mathbf{r}_{\text{ligand}}) + \Psi(\mathbf{r}_{\text{protein}}) - C,$$

where

$$C = 1/2[\int \Phi(\mathbf{r}_{\text{ligand}})\hat{p}(\mathbf{r}_{\text{ligand}})d\mathbf{r}_{\text{ligand}} + \int \Psi(\mathbf{r}_{\text{protein}})\hat{p}(\mathbf{r}_{\text{protein}})d\mathbf{r}_{\text{protein}}].$$

The constant C corrects the fact that interaction energies $H_{\text{lig.-prot.}}$ contribute to both the ligand molecular field $\Phi(\mathbf{r}_{\text{ligand}})$ and the protein molecular field $\Psi(\mathbf{r}_{\text{protein}})$.

The ligand $\hat{p}(\mathbf{r}_{\text{ligand}})$ and protein $\hat{p}(\mathbf{r}_{\text{protein}})$ probability distributions form a set of self-consistent mean field equations. In this study, there are S ligand and S protein conformations, and the probability distributions have the form

$$\hat{p}(\mathbf{r}_{\text{ligand}}) = \sum_{i=1}^{S} w^i_{\text{ligand}}\delta(\mathbf{r}_{\text{ligand}} - \mathbf{r}^i_{\text{ligand}})$$

and

$$\hat{p}(\mathbf{r}_{\text{protein}}) = \sum_{i=1}^{S} w^i_{\text{protein}}\delta(\mathbf{r}_{\text{protein}} - \mathbf{r}^i_{\text{protein}}),$$

where $\mathbf{r}^i_{\text{ligand}}$ and $\mathbf{r}^i_{\text{protein}}$ are the coordinates of ligand conformer i and protein conformer i. The weighting factors, w^i_{ligand} and w^i_{protein}, for the ligand and protein probability distributions were determined using an iterative algorithm (38, 41, 43). The resulting probability distribution $\hat{p}(\mathbf{r})$, which is the self-consistent solution to the mean field equations, corresponds to thermodynamic equilibrium of the ligand–protein system with the original Hamiltonian (Eq. **1**) that includes explicit ligand–protein interactions.

Results and Discussion

We report results of the mean field model applied to the structural assessment of HIV-1 complexes with inhibitors U85548e (47) and U75875 (48) (Fig. 1). We also hypothesize that the experimental difference in receptor-specific binding of these ligands to HIV-1 protease relative to its mutant, HIV-2 protease, can be rationalized on the basis of the energy spectrum analysis. The latter protein, which has mutations in the active site, has the same topology of the protein fold as HIV-1 protease with structural differences typical for protein fluctuations in the data base of HIV-1 protease complexes.

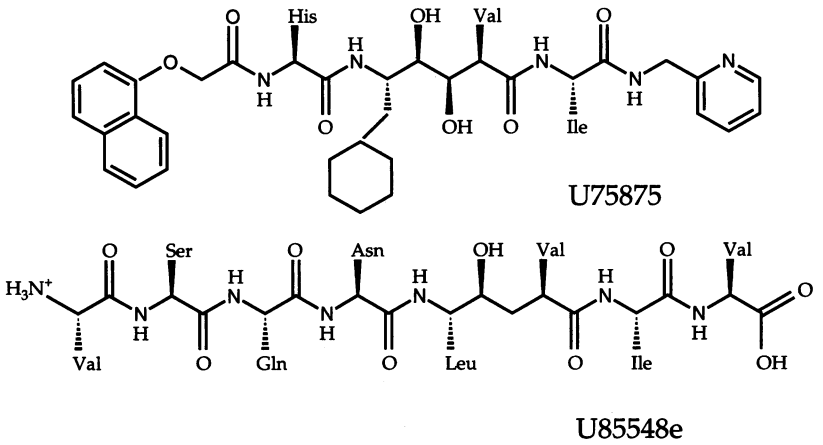

FIG. 1. Two-dimensional representation of the inhibitors U75875 and U85548e studied in this work. U75875 contains 45 hydrophobic and 14 hydrophilic atoms for a total of 59 heavy atoms. U85548e contains 40 hydrophobic and 21 hydrophilic atoms for a total of 61 heavy atoms.

These two proteases differ in their binding specificity for U75875, which has a binding constant to the HIV-2 protease of only 30 nM (49, 50) and <1 nM to the HIV-1 protease (48). In contrast, the proteins are less specific to the U85548e inhibitor, with a binding constant to the HIV-2 protease of 9 nM (51) and <1 nM to the HIV-1 protease (47).

We analyzed the thermodynamic aspects of molecular recognition and the relationship between receptor-specific binding of these ligands and their binding energy landscapes in HIV-1 protease. The molecular recognition model used in this study involves steric and hydrogen bond contributions to the ligand–protein interaction energy calculated from a piecewise linear potential summed over all protein and ligand nonhydrogen atoms (Fig. 2). There are four different atom types: hydrogen-bond donor, hydrogen-bond acceptor, both donor and acceptor, and nonpolar. Primary and secondary amines are defined to be donors, while oxygen and nitrogen atoms with no bound hydrogens are defined to be acceptors. Crystallographic water molecules and hydroxyl groups are defined to be both donor and acceptor, and carbon atoms are defined to be nonpolar. Each ligand–protein atom pair interacts through either the steric or the hydrogen bond potential (Table 1). The parameters (Table 2) were refined to yield the experimental crystallographic structure of a set of ligand–protein complexes as the global energy minimum (45).

The self-consistent iteration procedure (38, 41, 43) converges rapidly to the native complex for the U75875 inhibitor, independent of the initial probability distributions. For the U85548e inhibitor, depending on the initial distribution, the free-energy optimization converges to different metastable low-energy solutions. To analyze the thermodynamic implications of the results, we computed the entropy, the mean ligand–protein interaction energy, and the protein-order parameter as a function of temperature. The entropy is given as $S = \ln \Omega(E)$, where $\Omega(E)$ is the number density of states at the energy E. We define the continuum of the energy spectrum to be the region where $S > 0$ and the stability gap to be the difference between the ground state and the onset of this continuum. The temperature dependence of the protein order parameter (33), defined as $X(T) = 1 - \Sigma_{i=1}^{S} (w_{\text{protein}}^{i})^2$, monitors the transition from a multitude of protein conformations occupied by the complex at high temperatures to the native structure that becomes thermodynamically stable below the transition temperature. When the complex is locked in the ground-state conformer, the protein order parameter approaches zero, while at high temperatures the order parameter approaches unity.

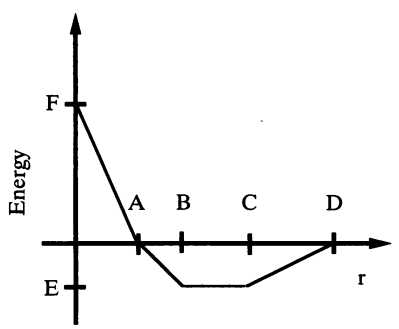

FIG. 2. The pairwise potential energy function $v_i^j(\mathbf{r})$ used to compute

$$H_{\text{lig.-prot.}}(\mathbf{r}_{\text{ligand}}, \mathbf{r}_{\text{protein}}) = \sum_{i=1}^{N_{\text{ligand}}} \sum_{j=1}^{N_{\text{protein}}} v_i^j(|\mathbf{r}_{\text{ligand}}^i - \mathbf{r}_{\text{protein}}^j|)$$

contains six parameters that depend on the interaction type. Hydrogen atoms are not included in the calculation. The units of energy are arbitrary.

We found that the energy spectra (Fig. 3), entropies (Fig. 4), mean energies (Fig. 5), and protein-order parameters (Fig. 6) contain important differences for the studied HIV-1 complexes. The global energy minimum for the U75875 inhibitor corresponds to the native complex and is separated by a definite gap from the quasicontinuous part of the spectrum. For the U85548e inhibitor, the low-energy part of the spectrum is dense, and the native complex corresponds only to one of the low-energy levels (Fig. 3). By contrast, inspection of the entropies, $S(E)$, reveals that the continuous part of both energy spectra is nearly identical (Fig. 4), as may be expected for inhibitors of similar size and composition (Fig. 1) (29, 52). If the temperature is high, there is no strong driving force that favors the native state for the U75875 ligand or any of the metastable solutions for the U85548e inhibitor. At high temperature, random ligand–protein conformations from the continuous part of the spectra dominate the thermodynamic equilibrium for both systems. As temperature decreases below

Table 1. Pairwise atomic interaction types for the molecular recognition model

Ligand atom type	Protein atom type			
	Donor	Acceptor	Both	Nonpolar
Donor	Steric	Hydrogen bond	Hydrogen bond	Steric
Acceptor	Hydrogen bond	Steric	Hydrogen bond	Steric
Both	Hydrogen bond	Hydrogen bond	Hydrogen bond	Steric
Nonpolar	Steric	Steric	Steric	Steric

Table 2. Parameters of the atomic pairwise ligand–protein potentials

Interaction type	A	B	C	D	E	F
Steric	3.4	3.6	4.5	5.5	−0.4	20.0
Hydrogen bond	2.3	2.6	3.1	3.4	−2.0	20.0

*A, B, C, and D are in Å. E and F are in arbitrary energy units.

the transition threshold, the native complex for the U75875 inhibitor becomes thermodynamically favorable, while the U85548e complex would "freeze" into one of the low-energy, metastable solutions. In fact, the mean field iteration procedure does not converge to a unique solution for the U85548e complex at low temperature but rather is sensitive to the initial conditions (Fig. 5). This signature of frustration is a consequence of the underlying energy spectrum and suggests that, while the continuous region of the energy spectrum is ligand independent, the low energy part of the spectrum describes the ligand-specific details of the protein response, analogous to nonself-averaging features detected in spin glass (23) and lattice protein-folding studies (29, 33). The temperature dependence of the mean interaction energy and the protein-order parameter (Figs. 5 and 6) has a pronounced sigmoidal shape for the U75875 inhibitor, which is an attribute of a cooperative transition, while for the U85548e ligand this shape is similar to that observed for random heteropolymers (23, 29, 33). Thus, the transition to the native complex for the U75875 ligand is sharp, as in a first-order phase transition. The difference in the curves for the two inhibitors is due to the distinct nature of the low-energy part of their energy spectra.

We have found that the energy landscape for the U75875 inhibitor has little frustration—i.e., the energy consistently decreases the more the structure of this complex resembles the native conformation (23, 24, 32). When the energy landscape is frustrated by a number of structurally distinct yet energetically similar levels, as with the U85548e inhibitor, the native complex would not dominate the equilibrium at temperatures relevant for experimental thermodynamic measurements. Hence, there is a complicated thermodynamic distribution for

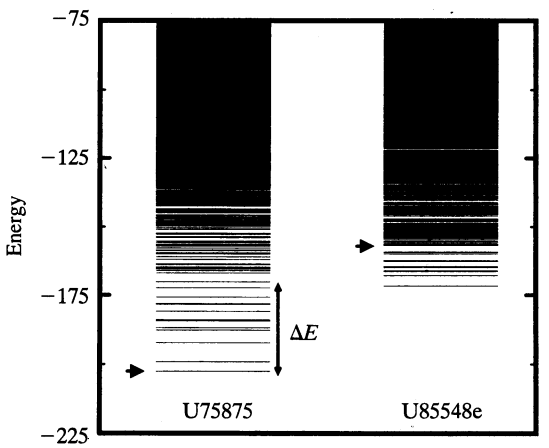

FIG. 3. The energy spectra for the U75875 and U85548e complexes with HIV-1 protease. In this study, there are $S * S$ states of the ligand–protein complex, where $S = 54$ is the number of distinct crystallographic protein conformers used to formulate the model. Each line represents the energy of a ligand–protein complex. The arrow to the left of each spectrum represents the energy of the native complex. The stability gap, ΔE, is shown to the right of the spectrum for the U75875 complex. For U85548e, there is no stability gap because the onset of the continuum of energy is coincident with the ground-state energy. Note that for U85548e the native-state energy is higher than the ground-state energy, while for U75875, the native state and the ground state are equivalent.

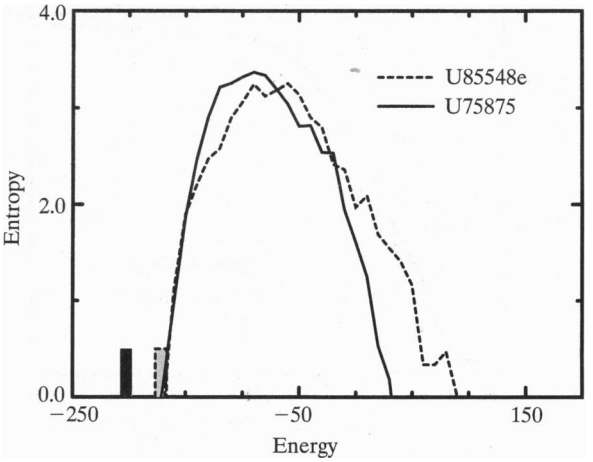

FIG. 4. The entropy of the U75875 and U85548e complexes as a function of energy. The number density was coarse-grained by using a bin width of 10 energy units. The curves were truncated at a value of $S = 0$ to facilitate viewing the onset of the continuum of energy states. The ground-state energy of each of the two inhibitors is shown on the plot to indicate the size of the stability gap. The ground state of U75875 is marked with a filled box, and the ground state of U85548e is marked with a shaded box surrounded by a dashed line. The onset of the continuum in both cases occurs at approximately $E = -170$.

the HIV-1 complex with the U85548e inhibitor even at low temperatures. Since energy levels in the continuum correspond to random ligand–protein arrangements with different HIV-1 protease conformers, they could serve as a measure of the interaction between the ligand and a protease with active-site mutations, such as HIV-2 protease. Therefore, the presence of a stability gap that separates the native U75875 complex from random arrangements, along with the lack of such a gap for the U85548e complex, provides a plausible rationale for the higher receptor specificity of U75875 to HIV-1 protease relative to U85548e. Although tight binding does not necessarily imply

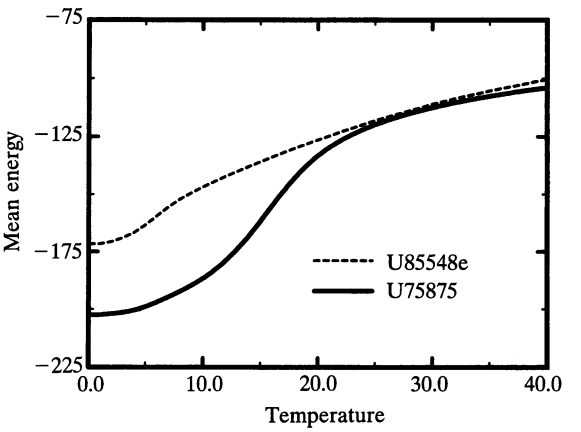

FIG. 5. Temperature dependence of the mean energies for the U75875 and U85548e complexes. The units of temperature are given in the same arbitrary units as the energy. To test sensitivity to initial conditions, three different sets of initial weights were chosen. First, the probabilities were set uniformly to $w^i_{\text{ligand}} = w^i_{\text{protein}} = 1/S$. Next, the probabilities were set to

$$w^i_{\text{ligand}} = w^i_{\text{protein}} = \exp[-H_{\text{lig.-prot.}}(\mathbf{r}^i_{\text{ligand}}, \mathbf{r}^i_{\text{protein}})]/ \sum_{i=1}^{S} \exp[-H_{\text{lig.-prot.}}(\mathbf{r}^i_{\text{ligand}}, \mathbf{r}^i_{\text{protein}})].$$

Finally, the initial probabilities were chosen to be a delta function centered at the native complex. For U85548e, the mean energy at low temperature is dependent on the initial conditions; only the curve that corresponds to the global free-energy minimum is shown.

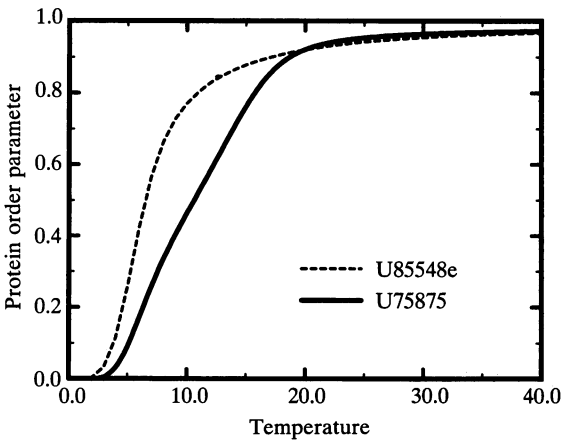

FIG. 6. Temperature dependence of the protein order parameters for the U75875 and U85548e complexes. For U85548e, the protein order parameter at low temperature is dependent on the initial conditions; only the curve that corresponds to the global free-energy minimum is shown.

the presence of a stability gap between the native state and the continuum, a minimally frustrated energy surface with a unique and stable native complex could be an important prerequisite for receptor-specific binding.

These results may have implications for site-specific ligand–protein and protein–DNA binding. On the basis of recent structural and thermodynamic experiments, it has been concluded that local folding events and disorder–order transitions could couple with binding (53). The local regions of HIV-1 protease (1–3), avidin (54), streptavidin (55), and *trp* repressor (56), partially disordered in the absence of a ligand, undergo folding transitions to create an ordered ligand–protein interface in the complex. The formation of avidin–biotin and streptavidin–biotin complexes leads not only to superstable, but also specific complexes (57). A mean field analysis of the energy spectrum may provide further insight into the effects of specificity for this class of ligand–protein complexes.

G.V. is grateful to Prof. A. Finkelstein and Dr. B. Reva for discussions on mean field theory and for making ref. 43 available prior to publication.

1. Wlodawer, A. & Erickson, J. W. (1993) *Annu. Rev. Biochem.* **62,** 543–585.
2. Abdel-Meguid, S. S. (1993) *Med. Res. Rev.* **13,** 731–778.
3. Appelt, K. (1993) *Perspect. Drug Discovery Design* **1,** 23–48.
4. Kuntz, I. D. (1992) *Science* **257,** 1078–1082.
5. Horton, N. & Lewis, M. (1992) *Protein Sci.* **1,** 169–181.
6. Miyamoto, S. & Kollman, P. A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8402–8406.
7. Krystek, S., Stouch, T. & Novotny, J. (1993) *J. Mol. Biol.* **234,** 661–679.
8. Wodak, S. J. & Janin, J. (1978) *J. Mol. Biol.* **124,** 323–342.
9. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R. & Ferrin, T. E. (1982) *J. Mol. Biol.* **161,** 269–288.
10. Goodsell, A. S. & Olson, A. J. (1990) *Proteins Struct. Funct. Genet.* **8,** 195–202.
11. Yue, S. Y. (1990) *Proteins Struct. Funct. Genet.* **4,** 177–184.
12. Shoichet, B. K. & Kuntz, I. D. (1991) *J. Mol. Biol.* **221,** 327–346.
13. Hart, T. N. & Read, R. J. (1992) *Proteins Struct. Funct. Genet.* **13,** 206–222.
14. Cherfils, J. & Janin, J. (1993) *Curr. Opin. Struct. Biol.* **3,** 265–269.
15. Stoddard, B. L. & Koshland, D. E. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 1146–1153.
16. Leach, A. R. (1994) *J. Mol. Biol.* **235,** 345–356.
17. Mizutani, M. Y., Tomioka, N. & Itai, A. (1994) *J. Mol. Biol.* **243,** 310–326.
18. Totrov, M. & Abagyan, R. (1994) *Nat. Struct. Biol.* **1,** 259–263.
19. Zacharias, M., Luty, B. A., Davis, M. E. & McCammon, J. A. (1994) *J. Mol. Biol.* **238,** 455–465.
20. Levinthal, C. (1968) *J. Chim. Phys.* **65,** 44–45.
21. Finkelstein, A. V. & Reva, B. A. (1992) *Protein Eng.* **5,** 617–624.
22. Reva, B. A. & Finkelstein, A. V. (1992) *Protein Eng.* **5,** 625–628.
23. Bryngelson, J. D. & Wolynes, P. G. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7524–7528.
24. Bryngelson, J. D. & Wolynes, P. G. (1989) *J. Phys. Chem.* **93,** 6902–6915.
25. Yue, K. & Dill, K. A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 638–642.
26. Skolnick, J. & Kolinski, A. (1990) *Science* **250,** 1121–1125.
27. Shakhnovich, E. I. & Gutin, A. M. (1990) *Nature (London)* **346,** 773–775.
28. Shakhnovich, E. I. & Gutin, A. M. (1993) *Protein Eng.* **6,** 793–800.
29. Shakhnovich, E. I. & Gutin, A. M. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7195–7199.
30. Kolinski, A. & Skolnick, J. (1993) *J. Chem. Phys.* **98,** 7420–7433.
31. Hinds, D. A. & Levitt, M. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 2536–2540.
32. Frauenfelder, H. & Wolynes, P. G. (1994) *Physics Today* **47,** 58–64.
33. Sali, A., Shakhnovich, E. I. & Karplus, M. (1994) *J. Mol. Biol.* **235,** 1614–1636.
34. Chan, H. S. & Dill, K. A. (1994) *J. Chem. Phys.* **100,** 9238–9257.
35. Goldstein, R. A., Lutney-Schulten, Z. A. & Wolynes, P. G. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9029–9033.
36. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995) *Proteins Struct. Funct. Genet.* **21,** 167–195.
37. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. & Socci, N. D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3626–3630.
38. Finkelstein, A. V. & Reva, B. A. (1991) *Nature (London)* **351,** 497–499.
39. Olszewski, K. A., Piela, L. & Scheraga, H. A. (1992) *J. Phys. Chem.* **96,** 4672–4676.
40. Roitberg, A. & Elber, R. (1991) *J. Chem. Phys.* **95,** 9277–9287.
41. Koehl, P. & Delarue, M. (1994) *J. Mol. Biol.* **239,** 249–275.
42. Lee, C. (1994) *J. Mol. Biol.* **236,** 918–939.
43. Finkelstein, A. V. & Reva, B. A. (1995) *Protein Eng.*, in press.
44. Forster, D. (1975) *Hydrodynamic Fluctuations, Broken Symmetry and Correlation Functions* (Benjamin Cummings, Reading, MA).
45. Gehlhaar, D. K., Verkhivker, G. M., Rejto, P. A., Sherman, C. J., Fogel, D. B., Fogel, L. J. & Freer, S. T. (1995) *Chem. Biol.* **2,** 317–324.
46. Bowie, J. U. & Eisenberg, D. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 4436–4440.
47. Jaskolski, M., Tomasseli, A. G., Sawyer, T. K., Staples, D. G., Heinrickson, R. L., Schneider, J., Kent, S. B. H. & Wlodawer, A. (1991) *Biochemistry* **30,** 1600–1609.
48. Thanki, N., Rao, J. K. M., Foundling, S. I., Howe, W. J., Moon, J. B., Hui, J. O., Tomasseli, A. G., Heinrickson, R. L., Thaisrivongs, S. & Wlodawer, A. (1992) *Protein Sci.* **1,** 1061–1072.
49. Ashorn, P., McQuade, T. J., Thaisrivongs, S., Tomasselli, A. G., Tarpley, W. G. & Moss, B. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 7472–7476.
50. Mulichak, A. M., Hui, J. O., Tomasseli, A. G., Heinrickson, R. L., Curry, K. A., Tomich, C.-S., Thaisrivongs, S., Sawyer, T. K. & Watenpaugh, K. D. (1993) *J. Biol. Chem.* **268,** 13103–13109.
51. Tomasseli, A. G., Hui, J. O., Sawyer, T. K., Staples, D. J., Bannow, C., Reardon, I. M., Howe, W. J., DeCamp, D. L., Craik, C. S. & Heinrickson, R. L. (1990) *J. Biol. Chem.* **265,** 14675–14683.
52. Pande, V. S., Grossberg, A. Y. & Tanaka, T. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 12976–12979.
53. Spolar, R. S. & Record, M. T., Jr. (1994) *Science* **263,** 777–784.
54. Livhan, O., Bayer, E. A., Wilcheck, M. & Sussman, J. L. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 5076–5080.
55. Weber, P. C., Wendolski, J. J., Pantoliano, M. W. & Salemme, F. R. (1992) *J. Am. Chem. Soc.* **114,** 3197–3200.
56. Zhao, D., Arrowsmith, C. H., Jia, X. & Jardetzky, O. (1993) *J. Mol. Biol.* **229,** 735–746.
57. Green, N. M. (1990) *Methods Enzymol.* **184,** 51–67.